# Conclusions and Recommendations from the PATH Weight of Evidence Workshop

**September 8-10, 1998**
**Vancouver, BC Canada**

# Conclusions and Recommendations from the PATH Weight of Evidence Workshop

## September 8-10, 1998
## Vancouver, BC Canada


The PATH Scientific Review Panel:

Steve Carpenter, University of Wisconsin
Jeremy Collie, University of Rhode Island
Saul Saila, University of Rhode Island
Carl Walters, University of British Columbia

Edited by:

Calvin Peters, Ian Parnell, and Dave Marmorek (ESSA Technologies Ltd.)
Robin Gregory (ValueScope Research)
Thomas Eppel (Decision Insights)


October 1, 1998

Citation: **PATH Scientific Review Panel.** Conclusions and recommendations from the PATH Weight of Evidence Workshop, September 8-10, 1998, Vancouver, BC Canada.

# Table of Contents

# List of Tables



# List of Figures

# The PATH Weight of Evidence Process

The PATH Weight of Evidence Process was developed to provide scientific input to decisions regarding actions to restore endangered stocks of Snake River spring/summer chinook salmon[1]. These decisions must be made in the face of uncertainty[2], because limitations in data on past and present conditions, uncertainties about future conditions (e.g., climate), and differences in the interpretation of existing data create uncertainties about the response of salmon stocks to management actions and environmental conditions. However, the existence of uncertainties does not imply paralysis in decision-making. Decisions must be made with full consideration of uncertainties; otherwise the risks to stocks may be underestimated. Techniques such as decision analysis help to clarify these risks by quantifying the effects of uncertainties.

PATH has narrowed the range of uncertainties and described the effects of remaining ones through the following analyses:

1. PATH retrospective analyses (Marmorek *et al.* 1996; Marmorek and Peters 1996) have reduced uncertainties related to causes of past trends in stocks, and clarified their relative importance. However, the evidence is not sufficient to resolve all uncertainties.
2. The PATH Preliminary Decision Analysis Report (Marmorek and Peters 1998) has characterized the relative effects of the remaining uncertainties on projected outcomes of alternative hydrosystem actions.
3. The PATH Weight of Evidence Report and its associated written submissions has summarized the evidence and other information related to the seven key uncertainties that have the greatest effects on the outcomes of management actions. This Report was completed through an iterative process of internal review and technical workshops with PATH scientists representing various tribal, state, and federal agencies, independent PATH scientists, and the PATH facilitation team.

The PATH Weight of Evidence Workshop was held September 8-10 1998 in Vancouver BC Its purpose was to provide a defensible process for using expert judgement, based on the evidence compiled in the Weight of Evidence Report, to quantify the relative degree of belief in the seven key uncertainties, and to document the experts' reasoning as a basis for gaining further insights about the overall modeling effort. It is important to recognize that weighting alternative hypotheses does not remove uncertainty — only additional evidence can do that. Instead, the weighting process provides a way to quantitatively incorporate the best available scientific information into decisions that must be made now, before additional evidence is available. The weightings are not intended to replace research, monitoring, or experimental management actions that could, if carried out properly, produce data that would in time narrow these uncertainties further.

---

[1] Actions considered were A1 (Current management), A2 (maximize transportation of smolts), and A3 (natural river drawdown of four Snake River dams). For A3, we consider two scenarios — one where drawdown is implemented in three years, and one in which drawdown is implemented in eight years. These two scenarios allow for possible delays in drawdown associated with obtaining congressional funding, possible litigation, and other non-biological uncertainties.

[2] We use 'uncertainty' as a general term for a process or model parameter for which there are data or information gaps that prevent complete knowledge. We use the term 'hypothesis' to describe potential explanations or numerical estimates for a given uncertainty. For example, the effectiveness of the predator removal program in improving survival rate of juvenile salmon in reservoirs is considered an uncertainty. One hypothesis about predator removal effectiveness is that the removal program has no effect on the survival rate of juvenile salmon in reservoirs; an alternative hypothesis is that survival rates of juvenile salmon are improved by 25% because of the removal program.

Relative weights on the critical alternative hypotheses were assessed by the four members of the Scientific Review Panel (see list of Workshop Participants below). Membership in the PATH SRP was determined by Dr. Larry Barnthouse, the PATH SRP coordinator. Members were selected from a candidate list based on ability in the areas of population biology, fisheries management, conservation biology, statistical analysis and modeling; national visibility; and lack of any current connection to Columbia River agencies. The SRP members thus have a broad and deep body of experience in fisheries population dynamics and modeling, and are acknowledged experts in this field. They are also familiar with technical aspects of PATH analyses through their reviews of previous PATH products. The workshop was also attended by:

1. process facilitators from ESSA, who were responsible for running the workshop;
2. elicitation facilitators, who led the formal expert judgment elicitations and assisted in the structuring of the questions that were evaluated by the scientific experts;
3. the three independent PATH scientists, who provided answers to technical questions posed by SRP members; and
4. The PATH SRP coordinator.

PATH scientists associated with tribal, state, and federal agencies did not directly participate in the weighting workshop, but provided input to the process through the Weight of Evidence Report and associated submissions.

## List of Workshop Participants

### PATH Scientific Review Panel

Carl Walters, U. of British Columbia
Jeremy Collie, University of Rhode Island
Saul Saila, University of Rhode Island
Steve Carpenter, University of Wisconsin

### PATH Independent Scientists

Randall Peterman. Simon Fraser University
Lou Botsford, U. of California, Davis
Rick Deriso, Inter-American Tropical Tuna Commission

### Elicitation Facilitators

Robin Gregory, ValueScope Research
Thomas Eppel, Decision Insights

### Process Facilitators

David Marmorek, ESSA
Calvin Peters, ESSA
Ian Parnell, ESSA

### PATH SRP Coordinator

Larry Barnthouse, LWB Environmental

# Objectives of the Weight of Evidence Workshop and Structure of this Report

Members of the SRP were asked to provide the following four items:

1. Relative weights on alternative hypotheses for key uncertainties, based on the material compiled in the PATH Weight of Evidence Report and submissions. These weightings were obtained through a formal process designed to elicit information from experts (the process is described in more detail in Section 1.2 below).
2. Judgements on whether the models are adequate for decision-making (i.e., projecting responses to hydro actions under consideration — A1, A2, A3).
3. Recommendations on whether and how to incorporate new hypotheses into the PATH analyses. These hypotheses were proposed as part of the Weight of Evidence Process, but were proposed after the main PATH modeling effort was completed (new hypotheses are described in section 5 of the PATH Weight of Evidence Report). The SRP was asked to comment on the relative merits of these hypotheses, whether or not their effects were already captured by existing hypotheses, and how feasible it would be to model them within the current PATH modeling framework.
4. Guidance on how best to collect information necessary to reduce remaining uncertainties.

The structure of this report follows these four objectives. Section 1 describes the process and results of the elicitation process, the rationale on which the weights are based, and the implications of the weights on projected outcomes of alternative actions. Section 2 summarizes the Panel's assessments of the models for decision-making. Section 3 summarizes the SRP's discussion of the merits of the new hypotheses and how they could be incorporated into the PATH analyses. Section 4 summarizes the Panel's discussion of strategies for collecting needed additional information.

# Section 1. Relative Weights on Key Hypotheses

## 1.1 Key Uncertainties and Alternative Hypotheses

Seven key uncertainties were identified in sensitivity analyses completed for the Weight of Evidence Report. The SRP agreed that these seven uncertainties were an appropriate focus for the Weight of Evidence process. These uncertainties and alternative hypotheses are briefly described below (more detailed description is provided in Table 3-2 of the PATH Weight of Evidence Report). Because the first three are the most influential, more time was spent at the workshop on these uncertainties than on the last four in the list below.

1. *Passage / transportation assumptions* – uncertainty in direct survival of in-river fish, and the partitioning of in-river survival between dam and reservoir survival. Transportation assumptions refer to uncertainty in the relative survival of transported and non-transported fish after they have exited the migration corridor (i.e., below Bonneville Dam). Passage and transportation assumptions are linked because they both rely on estimates of the survival rate of juvenile salmon that go through (in-river fish) or are transported around the hydropower system.

   CRiSP – a model of juvenile salmon passing through the hydropower system; generally results in higher estimates of survival rate of in-river survival and higher post-Bonneville survival of transported fish. Assumes that these survival rates have generally improved since 1980.

   FLUSH – a model of juvenile salmon passing through the hydropower system; generally results in lower estimates of survival rate of in-river survival and lower post-Bonneville survival of transported fish. Assumes that these survival rates have remained relatively constant since the late 1970s.

2. *Extra mortality* – Extra mortality is any mortality occurring outside of the juvenile migration corridor that is not accounted for by either: 1) productivity parameters in spawner-recruit relationships; 2) estimates of direct mortality within the migration corridor (from passage models); or 3) for the delta model only, common year effects affecting both Snake River and Lower Columbia River stocks. Extra mortality can in theory occur either before or after the hydropower migration corridor.

   Hydro – extra mortality is strongly associated with the mortality of in-river fish incurred while passing through the hydropower system. This is also described as the “here to stay unless dams are removed” hypothesis.

   BKD – extra mortality is caused by disease and other factors that will continue into the future. This is also referred to as the “here to stay” extra mortality hypothesis.

   Regime shift – extra mortality follows a cyclical pattern associated with 60-year cycles in ocean conditions.

3. *Life-cycle model*– uncertainty in the extent to which Snake River and lower Columbia stocks share common mortality effects.

   Alpha – a life-cycle model in which Snake River and lower Columbia stocks experience different mortality effects

   Delta – a life-cycle model in which there is a component of mortality that is shared by Snake River and lower Columbia stocks.

4. *Length of Transition Period* – duration of period between completion of dam removal and establishment of equilibrium conditions in the drawndown section of the river (transition period), reflecting uncertainty in the physical and biological responses to drawdown (e.g., short-term response of predators, release of sediment).

   Transition period is 2 years

   Transition period is 10 years

5. *Historical Turbine/Bypass Mortality* – uncertainty in historical estimates of bypass and turbine mortality for some projects in some years prior to 1980.

   TURB 4 – Bypass and turbine mortality higher, strongly related to rate of descaling

   TURB 5 – Bypass and turbine mortality lower, less strongly related to rate of descaling

6. *Predator Removal Effectiveness* – uncertainty in the effect of the predator removal program (i.e., removal of squawfish for bounties) on future survival of salmon smolts in reservoirs.

   The predator removal program will result in a 0% increase in reservoir survival of smolts

   The predator removal program will result in a 25% increase in reservoir survival of smolts

7. *Juvenile survival rate once river has reached equilibrium conditions after drawdown* – uncertainty in the long-term physical and ecological effects of drawdown (e.g., change in density of predators).

   Equilibrated juvenile survival rate over the 4-project section affected by drawdown is 0.85

   Equilibrated juvenile survival rate over the 4-project section affected by drawdown is 0.96

## 1.2 Process Used to Assign Weights

SRP members were led by the facilitators through a process for eliciting weights from experts. An expert judgment process is a structured procedure for eliciting opinions about uncertain variables from selected experts and for documenting these beliefs for others to evaluate. Formal expert judgment procedures are useful if existing models are controversial or inadequate, if a technical question is both important and complex, if key data are unavailable, and if one can expect significant scrutiny regarding these judgments. Expert judgment processes have been implemented successfully in a wide variety of domains and have provided insights to support analytical procedures undertaken by industry, government, and academic decision makers.

The primary purpose of expert judgments is to capture and reflect the current state of knowledge about a particular issue in the form of formal probability assessments. Some type of judgment is always necessary, and a formal process that relies on quantitative judgments has some key advantages over less formal approaches. The advantages of using a formal expert elicitation process include improvements in accuracy and consistency of the judgments, a more structured and focused dialogue among experts, and more effective ways to document and communicate the logic and assumptions of a particular analysis. Potential disadvantages include the required allocation of resources and the potential unfamiliarity of substantive experts with expressing their opinions in a formal and quantitative framework.

The expert judgment process (following procedures established as part of the NUREG-1150 process in the U.S.) should include the following steps:

1. Formulation of the technical question that is to be answered;
2. Selection of the experts;
3. Training of the experts regarding process and judgmental biases;
4. Decomposition of the technical question and clear definition of the variables;
5. Elicitation of probability distributions from individual experts;
6. Aggregation and discussion of individual differences among experts; and
7. Documentation (from substantive and normative experts) and communication.

The PATH Weight of Evidence workshop focussed on steps 3-7 of this process, with some modifications to better suit the situation at hand. First, in most cases, the conduct of an expert judgment elicitation requires two separate workshops along with a prior meeting to organize the sessions. In the PATH case, an initial one-day meeting was held in late July to discuss the elicitation process and to define realistic products, followed by the three-day workshop in September. Although the abbreviated process was necessary because of severe time and budget constraints (which limited the availability of the SRP members), the extensive work that already had been completed (as summarized in the PATH WOE report) made such an abbreviated process possible. Second, the experts were knowledgeable in probability theory and statistics, so that the training given to them (Step 3) was significantly shortened. Third, in a typical expert-judgment elicitation, hypotheses would be decomposed to identify those variables that are most responsible for differences in judgments across experts. Whenever possible, these would be constructed as continuous variables rather than as discrete values. However, PATH modeling analyses have structured alternative hypotheses either as discrete variables (e.g., predator removal effectiveness is either 0% or 25%), or as alternative models (e.g., Alpha or delta life-cycle model). Therefore, the primary need was for a process that would capture the experts' degree of belief in the form of relative weights on alternative hypotheses.

Relative weights on each of the key hypotheses (and the justification for those weights) were elicited through a structured interview process with SRP members. Interviews were done individually, to prevent individual members from influencing the assessments of others. Each interview was conducted by one of the elicitation facilitators, with at least one of the PATH independent scientists or process facilitators present to answer the expert's technical questions and to assist in documentation of the experts' judgments. After the interviews, SRP members discussed their rationales for assigning the weights they did. SRP members could elect to change their weights at this point based on the rationale of other members, but only one minor change was made. Rationales for weighting assessments are summarized below in Section 1.3.

SRP members applied four criteria to the evidence presented in the Weight of Evidence Report to evaluate alternative hypotheses:

1) *The **clarity** of the hypothesis*

   The intent of this criterion is to assess whether the hypothesis as implemented in various models clearly represents the effects that are intended. The clarity criterion does not favor single factor hypotheses, but where multiple factor hypotheses are proposed they should be structured in a way that clearly separates the effects of different factors.

2) *The existence of a reasonable mechanism or set of mechanisms by which the hypothesis operates*

   The hypothesis must propose a reasonable mechanism by which a given stress is converted into a change in survival. There should be evidence from physiological studies or direct survival measurements in the field to clearly associate a proposed stress with a reduction in survival.

3) *The consistency with empirical evidence*

   Measures of stock performance should vary inversely with the magnitude of the stressor across contrasts in space and time. Various measures of stock performance should be examined.

   Comparison of models and hypotheses to historical data is not necessarily a strong test of how those models will perform in the future, because in most cases, the actions that we are evaluating will create conditions that are outside of the range of historical observations. However, comparisons with historical data can still provide useful information for qualitative assessments of how reasonable (or unreasonable) is it for nature to behave in a way that would make this hypothesis valid.

4) *The validity of the method of projecting the hypothesis into the future*

   We need to assess whether the assumptions implied in methods of projection are reasonable. For example, do the methods properly reflect the hypotheses and mechanisms they were intended to represent? Do the methods generate reasonable escapement estimates under scenarios which extend current operations into the future?

For criteria 2 and 3, we further defined a set of criteria for evaluating empirical evidence.

- i) Applicability: Is the evidence relevant to the hypothesis being evaluated (i.e., is it the right stock, monitored in the right place at the right time)?
- ii) Clarity: Is the evidence clear, and not contested or confounded by other information, or an absence of sufficient good quality measurements?
- iii) Rigor: Is the evidence: 1) well established, generally accepted, peer reviewed[3] empirical evidence from relevant experiments and observations; 2) strong evidence but not fully conclusive; 3) theoretical support with some evidence; or 4) speculation or conjecture?

## 1.3 Relative Weights Resulting from Elicitation Process

Weights on the alternative hypotheses are summarized in Table 1. Although weights were elicited from the SRP members independently, for the most part they are generally consistent. It is important to remember that these weights do not necessarily represent absolute degrees of belief in the hypotheses. Rather, they are relative weights that represent the relative degree of belief in one hypothesis vs. another, out of a set of two or three alternatives. In some cases, SRP members did not think that any of the alternative hypotheses for a given uncertainty were very likely. This was the case with the predator removal effectiveness uncertainty, the length of the transition period, and with the passage models. These cases are documented further in the Rationale section (1.3) below.

[3] Peer review is recognized as having its limitations with respect to scrutiny of models (Starr *et al.* 1998).

**Table 1.** Relative weights assigned to alternative hypotheses by SRP members.

| Key Uncertainty | Alternative Hypothesis | Relative Weights | | | |
|---|---|---|---|---|---|
| | | JC | SS | SC | CW |
| Passage/transportation Models | FLUSH | 0.7 | 0.75 | 0.9 | 0.65 |
| | CRiSP | 0.3 | 0.25 | 0.1 | 0.35 |
| Extra Mortality | BKD (here to stay) | 0.3 | 0.25 | 0.495 | 0.4 |
| | Hydro (here to stay unless dams go) | 0.6 | 0.60 | 0.495 | 0.4 |
| | Regime Shift | 0.1 | 0.15 | 0.010 | 0.2 |
| Life-cycle models | Alpha | 0 | 0.7 | 0 | 0.1 |
| | Delta | 1 | 0.3 | 1 | 0.9 |
| Length of transition period | 2 years | 0.6 | 0.33 | 0.2 | 0.5 |
| | 10 years | 0.4 | 0.67 | 0.8 | 0.5 |
| Historical turbine/bypass mortality | TURB 4 (Higher) | 0.6 | 0.4 | 0.4 | 0.5 |
| | TURB 5 (Lower) | 0.4 | 0.6 | 0.6 | 0.5 |
| Predator Removal Effectiveness | 0% | 0.7 | 1 | 0.8 | 0.9 |
| | 25% | 0.3 | 0 | 0.2 | 0.1 |
| Equilibrated Juvenile Survival Rate | 0.85 | 0.6 | 0.8 | 0.8 | 0.25 |
| | 0.96 | 0.4 | 0.2 | 0.2 | 0.75 |

## 1.4 Rationale

SRP members based their weighting assessments on the criteria described above. Three out of the four SRP members assigned a quantitative score to each criterion, with 1 being the best score (met the criterion fully) and 4 being the worst (did not meet the criterion). The scores for each criterion were then subjectively integrated into an overall weight for each hypothesis. The fourth SRP member used the criteria as a qualitative guide for assigning weights but did not apply quantitative scores. Criteria scores are summarized in Table 2; summaries of verbal descriptions of rationale from the workshop are provided below.

**Table 2.** Criteria Scores – Summary (1=highest score, 4=lowest score).

| ***Key Uncertainty*** | **Alternative Hypothesis** | **Clarity** | | | **Mechanism** | | | **Evidence** | | | **Projection** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | JC | SS | SC* | JC | SS | SC | JC | SS | SC | JC | SS | SC |
| Passage / transportation Models | FLUSH | 2 | 2 | 3 | 3 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| | CRiSP | 4 | 3 | 4 | 4 | 2 | 3 | 3 | 3 | 3 | 4 | 2 | 4 |
| Extra Mortality | BKD | 3 | 3 | 1 | 4 | 3 | 1 | 4 | 2 | 2 | 3 | 3 | 2 |
| | Hydro | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 3 | 2 | 2 |
| | Regime Shift | 2 | 2 | 1 | 3 | 4 | 3 | 4 | 4 | 3 | 4 | 3 | 4 |
| Life-cycle models | Alpha | 4 | 2 | 1 | 4 | 1 | 1 | 4 | 3 | 4 | 4 | 2 | 1 |
| | Delta | 3 | 2 | 1 | 2 | 2 | 1 | 2 | 3 | 1 | 2 | 3 | 1 |
| Length of transition period | 2 years | 1 | 1 | | 2 | 2 | | 2 | 3 | | 3 | 3 | |
| | 10 years | 1 | 1 | | 3 | 3 | | 3 | 3 | | 3 | 3 | |
| Historical turbine /bypass mortality | TURB 4 (Higher) | 1 | 3 | | 3 | 3 | | 3 | 3 | | 3 | 3 | |
| | TURB 5 (Lower) | 3 | 2 | | 4 | 2 | | 3 | 2 | | 3 | 2 | |
| Predator Removal Effectiveness | 0% | 1 | 3 | | 2 | 1 | | 4 | 3 | | 2 | 3 | |
| | 25% | 1 | 3 | | 3 | 1 | | 4 | 3 | | 3 | 3 | |
| Equilibrated Juvenile Survival Rate | 0.85 | 1 | 2 | | 2 | 2 | | 2 | 3 | | 3 | 3 | |
| | 0.96 | 1 | 3 | | 2 | 2/3 | | 2 | 3 | | 3 | 3 | |

*Steve Carpenter applied the quantitative scoring system only for the three most important uncertainties.

## 1.5 Summaries of Rationales

### *1. Passage/Transportation Models*

**General Comments on Passage Models**

Some of the SRP members expressed the opinion that both passage models have deficiencies in empirical support, the level of complexity, etc. Therefore, the SRP was asked to comment specifically on the utility of the models for decision-making. These discussions are summarized in Section 2.

**Jeremy Collie (FLUSH = 0.7; CRiSP = 0.3)**

There are some problems with both models. Such detailed passage models are unnecessary for life-cycle modeling, and neither model can be fully validated. However, FLUSH is simpler and is based more on empirical relationships, while CRiSP is based on first principles. CRiSP is overparameterized and produces overly optimistic escapements when simulated with a policy resembling the status quo (A1). This is because CRiSP assumes higher $V_n$ and $D$ values prospectively.

**Saul Saila (FLUSH = 0.75; CRiSP = 0.25)**

The FLUSH model is a simpler, more robust model, requiring less manipulation of data than CRiSP. Also, FLUSH modelers have stronger ties to field work and empirical data collection than CRiSP modelers. The underlying survival vs. flow relationship in the FLUSH model has a stronger empirical basis, based on evidence from Atlantic salmon (Virtanen 1988; Hvidsten *et al.* 1995) and regression analyses of survival vs. flow rates (Cada *et al.* 1997; Saila and Ferson (in press)). Travel time is considered to be an especially important variable because of its effect on the physiology of juvenile salmon: as travel time increases, the fish are physiologically less able to swim and get safely to the ocean. CRiSP counter-arguments to the FLUSH mechanism in Table 4-4 were unconvincing. Projection of spawners from the FLUSH model are more conservative than those of CRiSP, and retrospective spawner estimates from FLUSH match the data better than CRiSP estimates, based on the BIC criterion (BIC scores are more relevant than AIC scores because they place greater importance on the number of parameters).

Virtanen, E. 1988. Smolting and osmoregulation of Baltic salmon, *Salmo salar* L., in fresh and brackish water. Finnish Fisheries Research 7:38-65.

Hvidsten, N.A. *et al*. 1995. Downstream migration of Atlantic salmon smolts in relation to water flow, water temperature, moon phase, and social interaction. Nordic Journal of Freshwater Research 70:38-48.

Cada, G.F., M.D. Duncan, S.V. Mitz, and M.S. Bavelhimer. 1997. Effects of water velocity on the survival of downstream-migrating juvenile salmon and steelhead: A review with emphasis on the Columbia River Basin. Reviews in Fisheries Science 5(2):131-183.

Saila, S.B. and S. Ferson (in press). Fuzzy regression in fisheries science: Some methods and applications. In: Fishery Stock Assessment Models. Alaska Sea Grant College Program: AK-SG-98-01, 1998:1-16.

**Steve Carpenter (FLUSH = 0.9; CRiSP = 0.1)**

Although FLUSH is a complex model (perhaps overly so), it gets a higher weight because it is simpler than CRiSP . One key distinction between the two models was in terms of the underlying mechanism, where the general rationale underlying the FLUSH model was felt to be significantly stronger. The empirical information also was felt to favor the FLUSH model by a wide margin. CRiSP is too complex, and the projections are overly optimistic. In particular, the high, constant D values in CRiSP are not credible — there is some risk of negative effects compounding over time because the effects of transport crowding and stress would result in adverse physiological and endocrine effects on fish.

**Carl Walters (FLUSH = 0.65; CRiSP = 0.35)**

FLUSH gets a higher weight because it is more clearly explained and is more parsimonious — it has fewer, but broader, mechanisms than CRiSP. The CRiSP model incorporates too many mechanisms, is overparameterized, and produces results that are overly optimistic. In addition, CRiSP's denial of the basic mechanism in FLUSH (an increasing mortality rate per time) is surprising. Empirical evidence for both models is low — partial reach survivals are matched by both models, but full reach survivals are not available because survival to below Bonneville is not measured. The FLUSH model alone explains historical declines in SARs. CRiSP needs a "demon in the ocean" to explain decline in SARs, but declines in SARs of other stocks that live in the same place in the ocean are not as severe and start later than declines in the Columbia. However, CRiSP should still get a reasonable overall probability because it can explain some empirical data (e.g., gas disease) better than FLUSH.

## *2. Extra Mortality*

**Jeremy Collie (BKD (here to stay) = 0.3; Hydro (here to stay unless dams removed) = 0.6; Regime Shift = 0.1)**

BKD represents irreversible changes in survival due to unspecified causes. Irreversible changes could also occur because of indirect genetic selection during barging, etc. There is a clear mechanism for reversible hydro effects. The weight on the hydro hypothesis is higher than BKD because it is grounded in the retrospective MLE analysis (Chapter 5 of Retrospective Analyses) — terms intended to represent the hydro effects were included in the spawner-recruit analyses. Hatchery and hydro effects are inseparable.

Regime shifts do occur but the timing and magnitude of the shift is not predictable. There is no evidence for a significant cycle with a 60-yr period. The time series that are referred to as evidence for the regime shift hypothesis show variability occurring at a range of frequencies. There is no empirical evidence that the regime shift impacts Snake River chinook salmon differently than other Columbia River salmon. Climate variations, resembling regime shifts, can and have been incorporated as autocorrelated year effects in the life-cycle model.

**Saul Saila (BKD = 0.25; Hydro = 0.60; Regime Shift = 0.15)**

The hydro hypothesis receives the highest weight based on the evidence presented in Table 4-15 of the Weight of Evidence Report (box a, counterpoints to box b), and because it is most consistent with historical patterns in recruitment- the problems facing salmonid survival were far less severe prior to construction of the dams. (Chapter 3 of the retrospective report). The evidence regarding BKD compiled by the State of Idaho (Submission 20 to the Weight of Evidence Report) was compelling, although some hatchery effects (disease, competition) are likely to continue even after dams are removed. Weight on regime shift is low because the empirical evidence for a regular, periodic 60-year cycle in environmental conditions is limited.

**Steve Carpenter (BKD = 0.495; Hydro = 0.495; Regime Shift = 0.01)**

All three of these alternative hypotheses were clearly stated, but the presentation of the arguments for the hypotheses in the PATH Weight of Evidence Report was not as clear. Of the three hypotheses, the hydro hypothesis and the BKD hypothesis are more consistent with empirical evidence, but differ in how they respond to changes in the hydrosystem. The hydro “here as long as the dams are” hypothesis has multiple mechanisms, including the damage inflicted on fish populations by passage through the dams as well as by adverse interactions between wild and hatchery stocks. BKD is just one of many mechanisms that would operate under a “here to stay” alternative; other factors which may be here to stay are disease, genetic changes, the presence of walleye, and changes in riparian habitat. There is no evidence to say whether the hydro response or the “here to stay” response is more likely. The regime shift hypothesis is vague and not well specified. The current implementation allows for an improvement in climatic conditions, but does not consider that things may get worse. If credence is given to a low-probability miracle, then at least an equal weight should be given to a low-probability catastrophe.

**Carl Walters (BKD = 0.4; Hydro = 0.4; Regime Shift = 0.1)**

Clarity of all extra mortality hypotheses is low with regard to which stocks and situations they apply to. The regime shift is particularly unclear. The main signal is in Alaskan stocks, and southern BC and Washington hatchery fall chinook stocks. Those who believe that the regime shift has occurred have been predicting a turn-around “next year” since the early 1980s.

Projection methods for the hydro and BKD hypothesis are straight-forward. A minor issue with implementing the regime shift is the lack of a stochastic element in the timing of the regime shift.

### *3. Life-Cycle Models*

**Jeremy Collie (Alpha = 0; Delta = 1)**

It should be possible to incorporate all hypotheses into a single model framework, although not all effects can be estimated simultaneously. The alpha model confounds anthropogenic and climatic effects. The alpha model assumes that the passage-model predictions are correct, then attempts to explain the residual with climate variables. However, the passage-model results are not independent variables, they are predicted values which vary with passage-model and TURB choices. There is a real possibility that the alpha model is explaining noise with the climate variables and it will always be possible to find a climate variable that correlates with noisy residuals.

The delta model was fit directly to spawner-recruit data and the model fits have been thoroughly scrutinized. Model diagnostics have not been provided or scrutinized for the alpha model.

**Saul Saila (Alpha = 0.7; Delta = 0.3)**

There are significant differences between chinook stocks that spawn in the Snake River system and those that spawn in lower Columbia systems. These differences are not reflected in common year effects. The greater distance and elevation that Snake River fish must traverse to spawn than the lower Columbia fish points to significant differences in physiological and endocrine systems. The Alpha model is a better portrayal of these differences than the Delta model. Further evidence for differences between the stock groups comes from Chapter 3 of the retrospective report — both groups of stocks have declined, but the Snake River fish have declined more.

Although the delta model explains historical data better than the alpha model (i.e., better fits to the SAR and spawner-recruit data, it may not be best for reasons which follow:

a) the future may not be described by the dame probability model as the past;
b) the model which includes the larger number of parameters may result in overfitting — which accounts for noise, not variability;
c) errors involved in fitting a larger number of parameters (including ratios — see Weight of Evidence Report Equation 4.2.2-7 vs. 4.2.2-1) may be more damaging to forecast accuracy, and
d) the preferred BIC goodness of fit statistic for the best alpha model is only 2% larger than the single BIC calculated for the delta model.

**Steve Carpenter (Alpha = 0; Delta = 1)**

The key to understanding this source of uncertainty is a close examination of the empirical evidence. In this case, the evidence for the delta hypothesis is believed to be much stronger than the Alpha model (AIC and BIC scores, and the fits to the SAR data, are better with the Delta model than with the Alpha model). This suggests that the alpha model should perhaps be omitted from further consideration. Additional analyses should be done to help further refine the delta model.

**Carl Walters (Alpha = 0.1; Delta = 0.9)**

There is a smooth decline in R/S residuals with distance for Snake River chinook and for other species. This argues against the Alpha model, which treats fish from the Snake River sub-region as a group. The Snake River basin is a diverse system — there is no basis for defining Snake River vs. non-Snake River stock groups.

### *4. Length of transition period*

**Jeremy Collie (2 yrs = 0.6; 10 yrs = 0.4)**

The shorter transition period receives a slightly higher weight because physical responses to drawdown are more important in determining spring/summer chinook juvenile survival rates than biological responses (juvenile spring/summer chinook are just passing through). If survival is related to flow and travel time, these may change relatively rapidly. The biological response is more hypothetical because there is less evidence to evaluate the timing of biological transitions. Predator populations could change rapidly due to movement and mortality in response to new flow conditions.

**Saul Saila (2 yrs = 0.33; 10 yrs = 0.67)**

Two years is very little time for salmon to adjust to any change in a natural system, so the 10 year hypothesis appeared more reasonable. The longer transition is also a more conservative estimate and is more consistent with a precautionary approach to fisheries management.

**Steve Carpenter (2 yrs = 0.2; 10 yrs = 0.8)**

This expert has extensive experience with food web manipulations in the Great Lakes region and believes that responses to changes in food webs in nature usually take longer to occur than we anticipate. Physical changes are likely to be longer than two years (Submission 13 to Weight of Evidence Report). Thus, the longer transition period is more reasonable.

**Carl Walters (2 yrs = 0.5; 10 yrs = 0.5)**

Initially, this expert placed a higher weight on the longer transition period. However, this was subsequently revised and he concluded that the evidence is insufficient to justify anything other than equal weights. Future modeling analyses should consider three options:

1. Immediate (next brood year);
2. Intermediate (5-8 years); and
3. Infinite (to allow for effects of exotic species).

### *5. Historical turbine/bypass mortality*

**Jeremy Collie (TURB4 = 0.6; TURB5 = 0.4)**

TURB1 has more straight-forward mortality estimates, but is apparently inappropriate for pre-1980 conditions. In general, there is a considerable amount of unresolved criticism with respect to both of these hypotheses (since both are dependent on assumptions about descaling). TURB 4 is believed to be more clearly defined since at least it is clearly dependent on the descaling mechanism, whereas TURB5 tries to "tweak" the descaling component and adds a second level of potential uncertainties. If mortality is not solely a function of descaling, why use a more convoluted function of descaling?

**Saul Saila (TURB4 = 0.4; TURB5 = 0.6)**

The limited amount of descaling information does not justify the emphasis on descaling in these hypotheses, so weights are close to 0.5. However, the weight on TURB 4 is slightly lower because the empirical basis for TURB4 is low, based on the BIC scores from Table 4-1 in the Weight of Evidence Report. TURB5 is a simpler and more logical hypothesis, and fits the historical spawner-recruit data better than TURB4. Also, TURB 5 is less optimistic than TURB 4 and is more consistent with a precautionary approach.

**Steve Carpenter (TURB4 = 0.4; TURB5 = 0.6)**

There is very little to distinguish between these hypotheses, and so weights of close to 5 and .5 are appropriate. However., TURB4 is the more optimistic hypothesis when applied into the future with the FLUSH model. In this case, it is wise to err on the side of caution, given the lack of sound empirical evidence relating to this judgment, and place a slightly higher weight on the more pessimistic TURB5 hypothesis.

**Carl Walters (TURB4 = 0.5; TURB5 = 0.5)**

Clarity and mechanisms for both hypotheses are low. Both are stories about what happened in those years, with no empirical evidence to support them.

### *6. Predator Removal Effectiveness*

**Jeremy Collie (0% = 0.7; 25% = 0.3)**

Both of these hypotheses are considered to be clearly defined. However, the mechanism for the 25% option is less clear because the relationship between the percentage of predators removed and the effectiveness of such a removal is not a linear one. The effects of predator removal depends on the predators being caught consistently over time and not being replaced with other predators. There are other predators of chinook salmon smolts than squawfish (e.g., terns).

**Saul Saila (0% = 1.0; 25% = 0.0)**

Predator removal may have short-term effects, but long-term benefits are unlikely because of changes in predator populations and communities. In fact, the occurrence of catastrophic predator/prey relationships in other systems argues that there is a low but non-negligible probability of a high consequence, adverse event occurring in the Columbia river system due to the actions of some predator (perhaps squawfish, perhaps another species). Therefore, the effects of predator removal may be either positive (improved survival of smolts) or negative (reduced survival). The higher weight on the no effect hypothesis reflects both possibilities.

**Steve Carpenter (0% = 0.8; 25% = 0.2)**

The predator removal argument hinges on the sustainability of predator suppression, leading to sustained reductions in predation on salmon. This is unlikely because predator suppression itself will create community dynamics that lead to oscillations in predation or demographic shifts in the predators (e.g. life history shifts to earlier age at first reproduction) that could actually increase total predation because of the allometry of bioenergetics. Furthermore, these reservoirs are open, highly disturbed systems susceptible to invasion by novel predators. Finally, the novel selection pressures on the salmon themselves (from dams, barging, hatcheries etc) may have reduced their capacity to adapt to novel predation regimes. All of these phenomena are known from other predator manipulations and could plausibly occur in the Snake/Columbia system. Of the two alternatives, the 0% option is closest to considering these effects, although this option is relatively optimistic. Certainly the highly optimistic +25% scenario should be balanced by an equally extreme pessimistic scenario.

**Carl Walters (0% = 0.9; 25% = 0.1)**

Clarity and mechanism for both hypotheses are low. Empirical evidence is low, and there are strong concerns about the evaluation methods of the predator removal studies in the Columbia River. Evidence from other systems in the Pacific Northwest (e.g., Alaska, Cultus Lake BC) is that while there may be short-term effects of predator removal programs, these initial effects do not persist because of compensatory responses in predator populations (e.g., only large squawfish are removed; smaller squawfish will continue to prey on salmon smolts), and community structure (squawfish replaced by some other predator species). Future modeling analyses should consider a negative benefit option.

### *7. Equilibrated Juvenile Survival Rate*

**Jeremy Collie (0.85 survival = 0.6; 0.96 survival = 0.4)**

It is unlikely that survival would return to the same levels estimated 30 years ago, given changes that have occurred in the river system (river course, exotics, etc.). Therefore, EJUV1 is slightly more likely.

**Saul Saila (0.85 survival = 0.8; 0.96 survival = 0.2)**

Current survival estimates are more applicable than the pre-dam estimates. The basic premise of EJUV1 (that there have been changes to the river that would prevent a return to pre-dam survival rates) is more reasonable. EJUV1 is also a more conservative estimate of the effects of drawdown on juvenile survival rates.

**Steve Carpenter (0.85 survival = 0.8; 0.96 survival = 0.2)**

The river has changed a great deal over the past two decades, so that more recent information (which suggests the lower equilibrated survival rate) would be more relevant.

**Carl Walters (0.85 survival = 0.25; 0.96 survival = 0.75)**

The higher survival rate receives more weight because extrapolating the lower equilibrated survival rate from the free-flowing reach to the entire reach leads to overall survival rates that are too low for a natural river (length of free-flowing reach is approximately $1/5^{th}$ of the length of the total reach; $0.85^5 = 0.44$).

## 1.6 Application of Weights to Modeling Results

### 1.6.1 Description of Performance Measures

The primary measures of performance used to evaluate the effects of management actions on survival and recovery of Snake River chinook salmon are the Jeopardy Standards, developed by the Biological Requirements Working Group (BRWG 1994) and largely accepted by NMFS (see Appendix D of the Preliminary Decision Analysis Report for details). There are two standards, the survival standard and the recovery standard. The Survival standard calculates the fraction of time during many simulations that the spawning abundance of a stock is above a certain specified low threshold. The threshold level used is either 150 spawners or 300 spawners depending on the characteristics of the stock and the stream. These levels were chosen because below these levels, spawner-recruit relationships are poorly known and unpredictable changes in population behavior are likely to occur. This fraction is calculated for simulations run over 24 years and simulations over 100 years. The survival standard is considered to be met when these fractions are 0.7 or higher.

The effect of a certain hydrosystem action on the chance of a spawning stock recovering is described by the Recovery standard chosen by the BRWG. The recovery standard calculates the fraction of simulation runs for which the average spawner abundance over the last 8 years of a 48-year simulation is greater than a specified level. The specified level of abundance (the recovery level) is different for each stream, and is 60% of the pre-1971 brood-year average spawner counts in each stream. An action is considered to have met the recovery standard when this fraction is 0.5 or greater.

For a particular action, each combination of hypotheses produces three outputs: the fraction of time during many simulations in which spawning escapement exceeds the survival threshold over 24 years, the fraction of time during many simulations in which spawning escapement exceeds the survival threshold over 100 years, and the fraction of simulation runs for which the average spawner abundance over the last 8 years of a 48-year simulation is greater than a specified level.

### 1.6.2 Results of Applying Weights to Modeling Results

The relative weights summarized in Table 1 can be applied to these model results to derive a weighted average[4] for each of the three outputs described above, for each action (Table 3; Figure 1). This analysis was done subsequent to the workshop, so that the implications of the weights were not presented or discussed at the workshop. The weighted averages show what the most likely outcome of the actions will be, given the uncertainties that affect future projections (although we emphasize that the range of possible outcomes is much greater than the range of most likely outcomes). We show the implications of each SRP member's set of weights separately, rather than combining the weights into an average or some other measure, because overall patterns in the results are similar across SRP members. General conclusions from Table 3 and Figure 1 are summarized below.

**General Conclusions:**

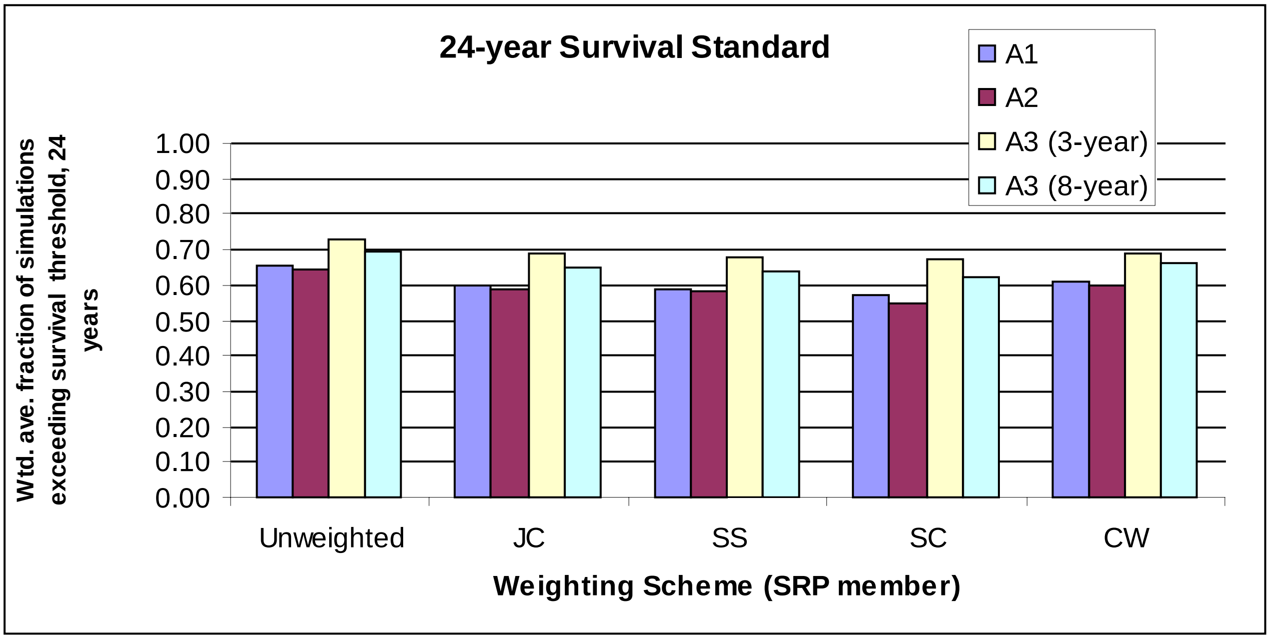
1. Outcomes for A3 are better than those of A1 and A2 for all jeopardy standards. The magnitude of the differences depend on the Jeopardy Standard and assumptions about when drawdown would be implemented. For the 24-year survival standard, differences in weighted average outcomes between A1 and A3 with a 3 year delay are around 0.1; with an eight year delay in implementing drawdown the differences are around 0.05. For the 100-year survival standard the differences are around 0.15, and for the 48-year recovery standard the differences are around 0.35.
2. None of the actions meet the 24-year survival standard (i.e., weighted average outcomes for all actions are less than 0.7). A1 and A2 are below the standard by around 0.1 – 0.15. Drawdown with a three-year delay comes within 0.01-0.03 of meeting this standard. Drawdown with an eight-year delay is below the standard by around 0.05.
3. Weighted average outcomes of A1 and A2 for the 100-year survival standard are right around the 0.7 standard. Under some weighting schemes, they meet the standard (maximum values are around 0.73), while under others they do not (minimum values are around 0.69). Results for A3 exceed the standard by around 0.17, regardless of whether a three-year or eight-year delay is assumed.

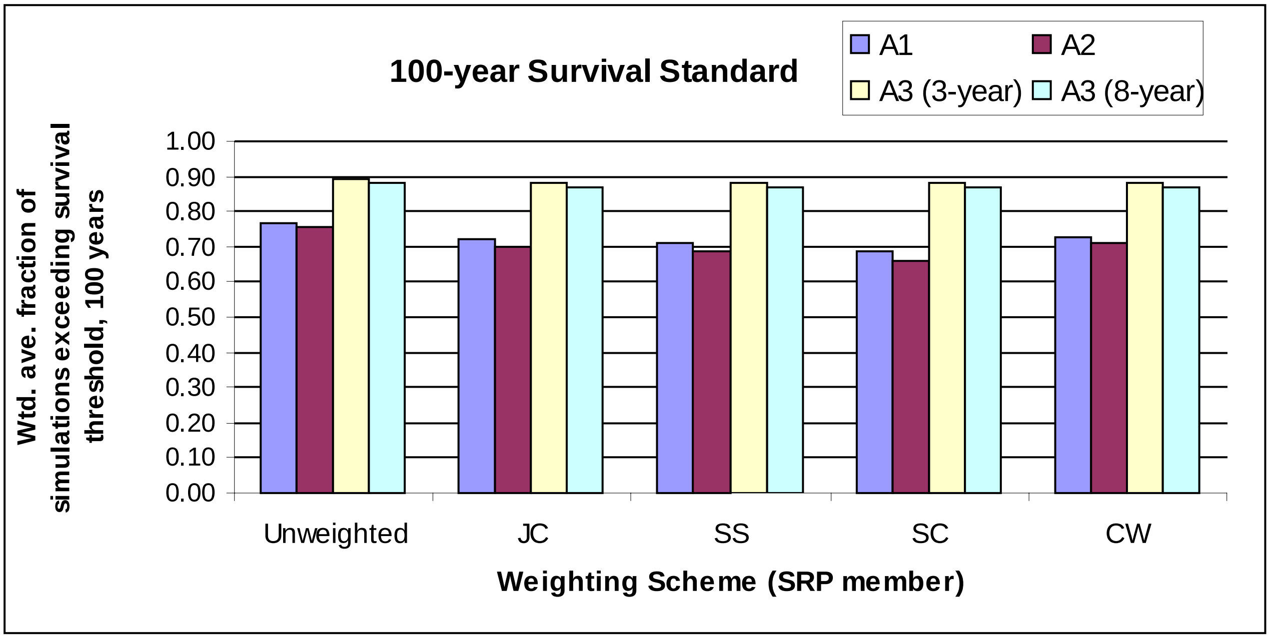
[4] The weighted average is calculated as follows: Each combination of hypotheses produces a set of three outputs (described above). The overall weight placed on a combination of hypotheses is the product of the weights assigned to the individual hypotheses that make up that combination. The weighted average outcome is calculated by multiplying the overall weight on each combination of hypotheses by the outcome for that particular combination, then summing over all combinations of hypotheses.

4. A1 and A2 do not meet the 48-year recovery standard; weighted average outcomes are less than the 0.5 standard by around 0.1. A3 exceeds the 0.5 standard by around 0.3, regardless of whether a three-year or eight-year delay is assumed.

**Table 3.** Weighted average outcomes of actions using weighting schemes developed by the SRP members.

| Weighting Scheme | Action | Weighted Average Performance Measure 24-year Survival | 100-year Survival | 48-year Recovery |
|---|---|---|---|---|
| JC | A1 | 0.60 | 0.72 | 0.41 |
| | A2 | 0.59 | 0.70 | 0.37 |
| | A3 (3 year) | 0.69 | 0.88 | 0.78 |
| | A3 (8 year) | 0.65 | 0.87 | 0.79 |
| SS | A1 | 0.59 | 0.71 | 0.40 |
| | A2 | 0.58 | 0.69 | 0.35 |
| | A3 (3 year) | 0.68 | 0.88 | 0.78 |
| | A3 (8 year) | 0.64 | 0.87 | 0.79 |
| SC | A1 | 0.57 | 0.69 | 0.35 |
| | A2 | 0.55 | 0.66 | 0.30 |
| | A3 (3 year) | 0.67 | 0.88 | 0.78 |
| | A3 (8 year) | 0.62 | 0.87 | 0.80 |
| CW | A1 | 0.61 | 0.73 | 0.43 |
| | A2 | 0.60 | 0.71 | 0.38 |
| | A3 (3 year) | 0.69 | 0.88 | 0.78 |
| | A3 (8 year) | 0.66 | 0.87 | 0.79 |





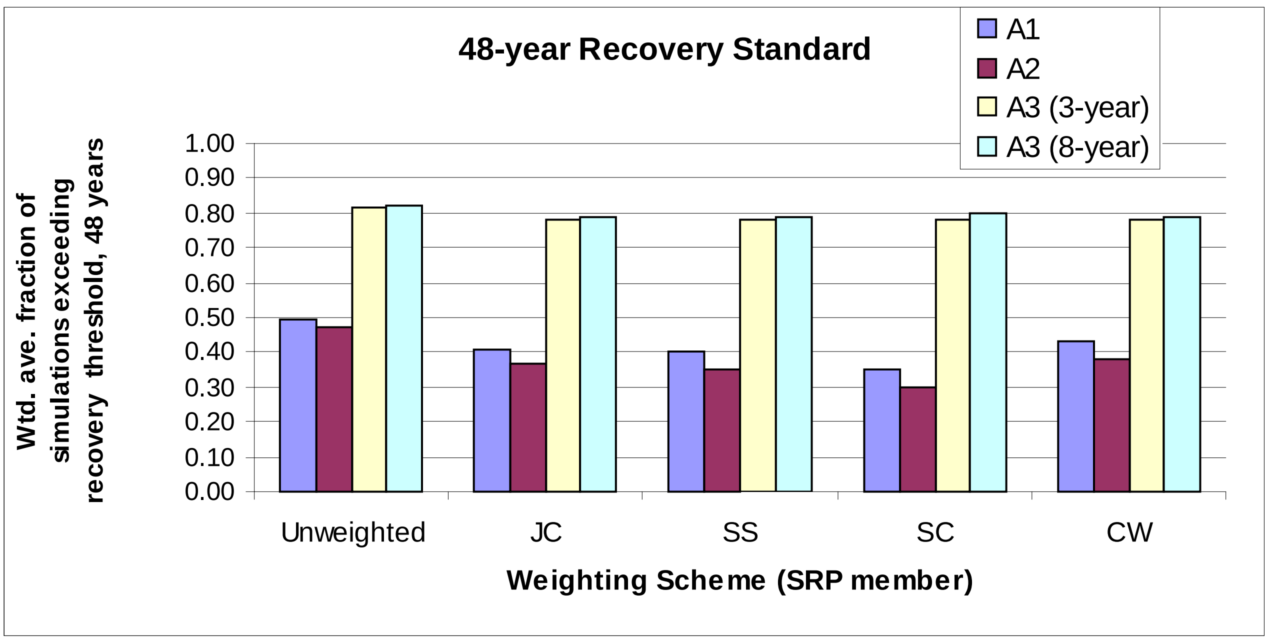




**Figure 1.** Weighted average outcomes of alternative actions, using weighting schemes assessed by the SRP members.

# Section 2. SRP Comments on the Utility of Using Models for Decision-Making

SRP members expressed concerns about both passage models in terms of their empirical support and level of complexity. Therefore, the SRP was asked to comment specifically on the adequacy of these models for PATH analyses.

The consensus of the Panel was that the adequacy of the models depended on the objective. The models are adequate for the limited objective of the current round of PATH analyses, which is to evaluate the relative benefits and risks to stocks of a limited set of hydrosystem actions (i.e., A1, A2, A3), under the current set of hypotheses. However, they may not be adequate for evaluating other actions or hypotheses, or for addressing objectives other than relative biological benefits/risks of actions (such as cost control, sustainability, learning). The models are also not adequate for addressing questions about mechanisms, and are too complex for assessing the relative merits of experimental management actions. The Panel made some suggestions on what type of model would be useful for assessing experimental management actions; these suggestions are included in Section 4.3 below.

# Section 3. SRP Comments on New Hypotheses

A number of new hypotheses were proposed after the main PATH modeling effort had been completed. Since PATH has not had time to model these hypotheses, the SRP was asked to comment on whether they should be pursued, based on their relative merits, the feasibility of implementing them, and the extent to which their effects had already been captured in existing hypotheses.

Eight new hypotheses have been proposed; these are described fully in Section 5 of the Weight of Evidence Report. Only four of these were specifically addressed by the SRP- the others were not felt to be developed sufficiently to warrant further examination. The four new hypotheses discussed by the SRP were:

1. Multi-factor extra mortality hypothesis (Weight of Evidence Report Section 5.1);
2. Hatchery extra mortality hypothesis (Weight of Evidence Report Section 5.2);
3. Reformulation of the hydro extra mortality hypothesis (Weight of Evidence Report Section 4.2.3); and
4. Juvenile survival rate during the transition phase (Weight of Evidence Report Section 5.4).

## 3.1 Multi-factor Extra Mortality Hypothesis

### Brief Description

Extra mortality of Snake River chinook salmon is a function of climatic effects, bird predation in the estuary, hatchery interactions, and hydrosystem effects.

### SRP's Evaluation

The SRP concluded that there was no need to pursue this hypothesis any further for the following reasons:

a) The hypothesis is based on the presumption that a small number of factors can explain mortality. However, identifying the complete list of these factors is impossible, and directs attention away from the fundamental uncertainties. The SRP strongly warned against confusing the major issues by introducing hypotheses that involve complex interactions between many factors.
b) The magnitude and timing of the effects of multiple factors cannot be estimated simultaneously, so this hypothesis is untestable. Factors are either confounded (e.g., hatcheries and hydrosystem effects) or are outside of the retrospective data (e.g., bird predation, which is assumed to have increased dramatically since 1990).
c) Most SRP members had already considered elements of the multi-factor hypothesis when assigning weights to the existing extra mortality hypotheses. For example, predation by birds was considered to be part of the BKD hypothesis (sampling from 1977-1992 extra mortality is appropriate because other predators may have been acting in the past as birds are now) and the hydro hypothesis (because of interactions with transportation). Hatchery effects were considered by many to be a component of the hydro hypothesis. Therefore, considering a multi-factor hypothesis is not likely to add anything to the existing analysis.

## 3.2 Hatchery Extra Mortality Hypothesis

### Brief Description

Releases of hatchery smolts (particularly steelhead smolts) from Snake River hatcheries have had a negative effect on survival of Snake River spring/summer chinook salmon.

### SRP's Evaluation

The SRP felt that it is reasonable to hypothesize that interactions with hatchery fish can affect the survival of wild fish. Figure 5 in Submission 1 to the Weight of Evidence Report provides the basis for a simple prediction of the effects of hatcheries on extra mortality. With the FLUSH model, extra mortality increased from around 0.4 to around 0.8 as production of hatchery smolts increased. With the CRiSP model, extra mortality increased from around 0.25 to 0.8.

However, the SRP concluded that it was not feasible to implement the hatchery hypothesis in the models at this late date because:

a) hatchery effects are confounded with development of the hydrosystem;
b) the distinction between the hatchery hypothesis and the other extra mortality hypotheses is not clear; and
c) the responses of fish to the different actions under these hypotheses, is also not clear.

Some SRP members considered hatchery effects as part of hydro effects (because of crowding in barges and forebays, etc.), and assumed that if dams were removed then the hatchery production would be reduced. Others assumed that the reversible component of hatchery effects was captured in the hydro hypothesis, while the irreversible component was captured in the BKD hypothesis.

The SRP suggested that hatchery effects and hydrosystem effects can only be separated through experimental manipulation. Hypothesized hatchery effects should be considered explicitly in the evaluation of an experimental reduction in hatchery production (see discussion of experimental management actions in Section 4).

## 3.3 Reformulation of the Hydro Extra Mortality Hypothesis

### Brief Description

The current formulation of the hydro extra mortality hypothesis assumes that the extra mortality of non-transported fish is proportional to the in-river survival of non-transported fish, with a different proportionality constant in each year. Some PATH members have questioned this assumption (Submission 3 to the Weight of Evidence Report). In response, an alternative formulation was proposed. Under this reformulation of the hydro hypothesis, extra mortality is essentially "here to stay" under actions A1 and A2 (i.e., is similar to the BKD hypothesis), and reverts to pre-dam levels when dams are removed under action A3.

### SRP Evaluation

Some members of the SRP felt that the new formulation was preferable to the current formulation because it was simpler. However, the results of using the reformulated hydro hypothesis are predictable without doing any additional model runs. Outcomes of actions A1 and A2 would be similar to those under the current BKD hypothesis, which are generally lower than outcomes under the current hydro hypothesis. Outcomes of action A3 would be slightly higher, based on results of sensitivity analyses presented in Appendix H of the Weight of Evidence Report. Therefore, implementing the reformulated hydro extra mortality hypothesis would tend to increase the difference between A3 and the other actions. Given that the weights on the hydro hypothesis are fairly high (ranging from 0.4 to 0.6; Table 1), one would expect to see a slight decrease in weighted average jeopardy outcomes under A1 and A2, and a slight increase in weighted average outcomes under A3 resulting from the reformulated hydro hypothesis. Because the 24-year survival standard outcomes for A3 are only slightly below the standard of 0.7, the increase resulting from the reformulated hydro hypothesis would increase the chances of A3 meeting that jeopardy standard.

## 3.4 Juvenile Survival Rate During the Transition Phase

### Brief Description

Drawdown of reservoirs will lead to an initial reduction in reservoir survival of smolts because predators will be concentrated into a reduced volume of water. Current drawdown scenarios assume that juvenile survival rates through the drawndown reach will increase linearly to equilibrated values.

### SRP Evaluation

The SRP concluded that this hypothesis did not need to be explicitly modeled, for the following reasons:

a) Sensitivity analyses presented in section 5.4 of the Weight of Evidence Report showed that even an initial 50% reduction in reservoir survival had virtually no effect on model results.

b) Most SRP members took predator responses into account when assigning weights to the existing hypotheses about transition period (i.e., placed higher weights on the longer transition period to account for delays in responses of predators) and equilibrated survival rate (i.e., placed higher weights on the lower equilibrated juvenile survival rate to account for long-term changes in predator responses).

# Section 4. SRP Comments on Experimental Management

In past reviews of PATH products, the SRP has commented repeatedly on the need for an experimental management approach to resolving major uncertainties. At this workshop, they were asked to provide additional comments in light of the evidence compiled in the Weight of Evidence Report. Their discussions are organized into the following sections:

1. Decision-making under uncertainty and the need for experimental management;
2. Options for experimental management actions; and
3. Evaluating experimental management options.

## 4.1 Decision-Making Under Uncertainty and the Need for Experimental Management

The weights assigned by SRP members to the key uncertainties reflect the relative likelihood of the alternative hypotheses, based on the evidence currently available. However, all SRP members commented that in some cases, the empirical evidence on which to evaluate alternative hypotheses was poor or lacking. This is because many events have occurred outside of the temporal and spatial range of historic monitoring programs, and outside of our experience. In the face of this level of uncertainty, the SRP felt that it is unrealistic and imprudent to expect irreversible, long-term decisions to recover stocks because there is little confidence that these actions will have the effects they are projected to have. **However, the SRP strongly cautioned that uncertainty should not be used to justify either delaying action or taking no action at all.** Such a misuse of uncertainty in decision-making is not an acceptable component of responsible fisheries management (United Nations Precautionary Approach). Instead, the SRP noted that the existence of uncertainties points to the need to take actions that:

a) result in the best chance at survival and recovery of stocks; and
b) generate information to reduce uncertainties and improve future decision-making.

Carefully designed and implemented experimental management actions provide that opportunity.

## 4.2 Options for Experimental Management

The SRP identified three major potential factors amenable to experimental management: hydrosystem effects, hatchery effects, and transportation effects. This suggested three possible experimental manipulations: dam removal, elimination or substantial reduction of hatchery releases, and transportation turn-off (Table 4). Implementing these actions in a well-designed experimental fashion can provide stronger evidence for the relative effects of each factor.

**Table 4.** Main experimental management options and combinations

| Dam Removal | Eliminate Hatchery Releases | Transportation turn-off |
|---|---|---|
| X | | |
| | X | |
| | | X |
| X | X | |
| | X | X |
| X | | X |
| X | X | X |

The SRP evaluated two strategies for implementing these actions. Both strategies are likely to require several decades to detect shifts in stock performance because of the high degree of natural variability in these measures.

1. Incremental alternative: implement the cheapest action first and monitor effects, then progressively more costly ones. This strategy requires the lowest up-front costs. However, it is the most risk-prone of the two strategies, because more effective actions are delayed when the cheapest action fails to produce the desired response in fish stocks. It may also be the higher-cost option in the long-term, if mitigative actions are required as each treatment is assessed. The incremental approach characterizes past management in the Columbia River.
2. "Reverse staircase" alternative: Implement all actions at once, then turn dams, hatcheries, or transportation back on one at a time. This is a more risk-averse approach than the incremental approach and is more likely to lead to stock recovery, but involves larger up-front costs. Long-term costs depend on the costs associated with reversing the experimental actions (i.e., replacing dams, turning hatcheries and transportation back on).

## 4.3 Evaluating Experimental Management Actions

Experimental management actions can be evaluated by assuming some underlying response, then simulating the effects of an action to see if they are able to detect the assumed response, and how long it would take before the response is detected. The SRP made the following suggestions for evaluating experimental management options.

Evaluation of experimental actions will require a much smaller and simpler set of hypotheses and models than those used in the current PATH analyses. The change in recruitment anomalies (residuals from spawner-recruit functions) is the crucial observation to distinguish among the hypotheses (Transport:control ratios should also be monitored for differential effects of hatcheries on transported vs. non-transported fish). Therefore, hypotheses should focus on the expected response of R/S to an experimental action. For example, the information presented in Figure 5 of Submission 1 to the Weight of Evidence Report could form the basis for a simple hypothesis about the effects of hatchery actions on extra mortality.

Models used in evaluating experimental management actions need only to be able to predict recruitment anomalies in response to the experimental actions, under different hypotheses about the responses of R/S to the experimental treatment, while removing the effects of other factors on R/S. Density-dependent spawner effects, common year effects, and the main in-river survival effects of lower dams should be factored out. The last of these components, in-river survival effects of lower dams, will require some form of passage model coupled with a life-cycle model. However, SRP members felt that the existing passage models were too complex, time-consuming, and inflexible for exploring experimental management options. They recommended that a simpler, more flexible passage model be used instead. This might be a simplified version of FLUSH and CRiSP (as proposed by Jim Anderson in his e-mail of Nov. 5 1997), or a simple proportional relationship between mu and WTT, as was tested in Chapter 5 of the Retrospective Report.

# Appendix A. Conclusions and Recommendations of the Elicitation Facilitators (Gregory and Eppel)

We believe that the PATH WOE workshop succeeded well in achieving its stated purpose of determining and discussing relative weights on the alternative hypotheses that were considered for each of the seven identified key sources of uncertainty. This information will be of assistance to PATH team members as they complete their analysis of model runs over the coming months, and it will be helpful to scientists and policy makers on the Implementation Team as they struggle with the difficult task of developing clear recommendations for management of the salmon resources of the Columbia River.

We also believe that substantial additional information on hypotheses and issues relating to salmonid survival was obtained in the course of the three-day workshop and that these issues bear closer scrutiny at some point in the PATH process, because of the insights they may provide and because of their importance to the presentation and communication of information from PATH members to the Implementation Team as well as to other State or Federal decision makers. Four of these issues are noted briefly in this concluding section.

*Model complexity*

All four of the experts whose judgments were elicited freely expressed their frustration with what is perceived to be an unduly high level of complexity in both the FLUSH and CRiSP models. Although everyone at the workshop agreed that the CRiSP model was the worst, even the FLUSH model was viewed as suspect because of its complexity. The experts noted several instances when improvements in information could not be incorporated effectively because of the length of time it would require to run the model. This issue was viewed as particularly important in the context of developing an experimentally based management system on the Columbia River that could, over time, lead to improvements in our knowledge of the system: by and large, participants agreed that the complexity of both the FLUSH and CRiSP models would render them next-to-useless as tools for testing adaptive management options.

*Adaptive management*

Strong support was expressed among the experts for some form of an adaptive management approach to fisheries policies on the Columbia. In comparing differences between policies currently under consideration (A1, A2, A3) and policies explicitly advocating postponement of a decision until learning-based experiments were completed (i.e., adaptive management), it was thought to be important to develop a strategy for expressing the benefits of an adaptive management approach in language that could be understood by the Implementation Team and by government decision makers. This choice is depicted in the decision tree shown on the next page (Figure A-1), which compares several possible actions and adaptive management scenarios. It was further noted that approaches involving the drawdown of dams could be viewed in stronger or weaker adaptive terms, based on the associated type and level of experimentation and on the plan for evaluation of results over time.

*Value of information*

As part of any adaptive management strategy, we strongly urge that explicit analyses be conducted to anticipate the value of the information about the key variables that might impact the choice of a preferred alternative. Value-of-information studies are a central concern of decision analysts and techniques have been developed for their conduct that could both clarify the expected value of alternative studies and help to communicate between diverse disciplines and decision makers involved in Columbia River resource management strategies.

*Linkage to other Columbia River studies*

The output of PATH is expected to be compelling to ecologists and fisheries biologists. We are concerned that it may not be as compelling as possible to others involved in Columbia River resource management decisions; in particular, that PATH analyses may not speak clearly to those parties concerned with developing economic and social aspects of the Columbia River spring/summer chinook salmon EIS. With this concern in mind, we advocate that an attempt be made to elicit the values of these interdisciplinary colleagues (i.e., what types and levels of impacts concern them the most?) so that their concerns can be addressed directly as part of PATH documentation. We also advocate that the documentation of PATH results be proactive, in the sense of introducing arguments that address some of the anticipated reasons why (for example) adaptive management strategies might tend to be disregarded (e.g., through the effects of a positive discount rate, which would tend to bias management strategies toward short-term benefits).

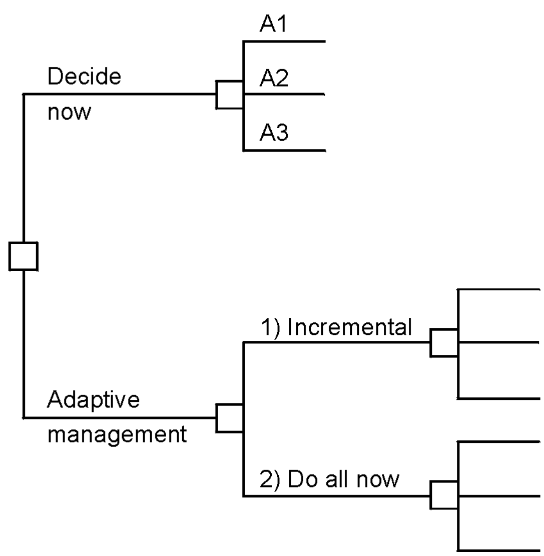



**Figure A-1.** Decision tree showing possible experimental management scenarios.